

# Monte Carlo Based Identification Strategies for State Space Models

Ing. Milan Papež

Posudek disertační práce

Disertační práce Ing. Milana Papeže se zabývá využitím bayesovského přístupu a jeho aproximací výpočetními metodami Monte Carlo při identifikaci dynamického systému a odhadu stavu s důrazem na efektivní numerickou implementaci.

## *Zhodnocení významu disertační práce pro obor*

Zvolené téma práce je aktuální, a to jak z hlediska rozvoje teoretického aparátu, tak z hlediska aplikací. Využití metod Monte Carlo významně rozšiřuje praktickou aplikovatelnost bayesovského přístupu na nelineární a/nebo negaussovské problémy. Aktuálnost tématu a intenzita výzkumu v této oblasti je dokumentovaná také kvalitním seznamem referencí, který obsahují více než 200 položek a dokládá, že doktorand má dobrý přehled o současném stavu poznání v této oblasti.

## *Vyjádření k postupu řešení problému, použitým metodám a splnění určeného cíle*

Postup řešení problému, použitý doktorandem, je logický a dobře strukturovaný. V kapitole 1 doktorand popisuje základní koncepty metod Monte Carlo a pokročilých metod jejich numericky efektivní implementace. V kapitole 2 je diskutována aplikace těchto metod na úlohy odhadu stavu a parametrů. Tyto kapitoly jsou přehledového rázu.

V dalších kapitolách doktorand formuluje specifické cíle práce a jejich řešení. Tyto cíle se dají shrnout do tří oblastí.

1. Kapitola 3 popisuje využití metod Monte Carlo pro odhad stavu a parametrů systémů s nelineárním vývojem stavu (nepoddajná část struktury modelu řešená aproximačními metodami) a lineárním modelem vývoje parametrů (poddajná část struktury modelu řešitelná analyticky). Přínosem autora je využití této struktury modelu k dosažení větší numerické efektivnosti a vyšší přesnosti algoritmů odhadu. Kapitola 4 rozšiřuje tento přístup o možnost sledování časově proměnných parametrů modelu aplikací zobecněných metod zapomínání na poddajnou část struktury modelu, kde je možné najít bayesovské řešení ve formě algebraické rekurze.

2. Kapitola 5 řeší problém odhadu stavu pro skokové Markovské procesy. Autor odvozuje metodu Rao-Blackwell filtrace stavu a poté v kapitole 6 metodu maximální věrohodnosti pro odhad parametrů. Tu pak kombinuje s výsledky kapitoly 5 přístupem EM (Expectation–Maximization).
3. Třetí okruh problémů je formulován a řešen v kapitole 7, která se zabývá strategií přenosu znalosti (“transfer learning”) mezi dvojicí Kalmanových filtrů bez dostupné informace o statistické závislosti jejich dat. Doktorand formuluje strategii založenou na konceptu „Fully Probabilistic Design“ a ukazuje, že tato metoda může výrazně zlepšit kvalitu přenosu znalosti, dokonce i ve srovnání s metodami, které informaci o statistické závislosti dat používají.

Použité metody založené na využití bayesovského přístupu a jeho Monte Carlo aproximací jsou vhodným nástrojem k dosažení deklarovaných cílů práce. Oceňuji, že autor detailně vysvětluje některé obecně přijímané postupy (např. “Assumed Density Approximation”, metodu „Expectation-Maximization“ nebo „Fully Probabilistic Design“) a je schopen tyto poměrně složité koncepty s dobrou matematickou erudicí aplikovat.

Lze konstatovat, že jednotlivé cíle stanovené v kapitolách 3 - 7 byly splněny a každá z těchto kapitol je uzavřena popisem experimentálních výsledků a jejich detailní diskusí. Za určitou slabinu práce považuji volbu modelů pro experimentální ověřování dosažených výsledků. Všechny modely (s výjimkou kapitoly 7) mají pouze skalární stavovou veličinu a nelineární modely jsou navíc velmi akademické – není proto možné usoudit, jaký přínos by prezentované výsledky měly pro řešení konkrétní reálných problémů.

#### ***Stanovisko k výsledkům disertační práce a k původnímu přínosu autora***

Presentované výsledky dokládají schopnost uchazeče samostatně vědecky pracovat a jsou původním přínosem autora. V rámci účasti na kvalitních mezinárodních konferencích prošly oponentním řízením a byly přijaty odbornou komunitou.

#### ***Vyjádření k systematické, přehlednosti, formální úpravě a jazykové úrovni práce***

Uspořádání práce je na první pohled systematické a logické, ale při bližším seznámení s prací se ukazuje, že není opravdu dobře promyšlené. Řada generických partií se v textu na mnoha místech opakuje. Práce je proto zbytečně rozsáhlá a její studium není efektivní.

Výklad některých částí je navíc obtížně srozumitelný. V kapitole 1 se pracuje s notací, která není v textu dostatečně vysvětlena (např. pro mnohonásobně přetížený symbol pro distribuci  $\pi(\cdot)$ ).

není jasně definováno, zda se jedná o hustotu pravděpodobnost nebo kumulativní distribuční funkci, pro výpočet střední hodnoty odpovídající těmto interpretacím  $E\{f\} = \int f(x)\pi(x)dx$  nebo  $E\{f\} = \int f(x)d\pi(x)$  se používá nestandardní zápis  $E\{f\} = \int f(x)\pi(dx)$ , vztahy typu (1.26) jsou pro nezasvěceného čtenáře rovněž obtížně nesrozumitelné, není formálně rozlišena  $t$ -rozměrná sdružená hustota pravděpodobnost veličin  $p(x_1, \dots, x_t) = p(x_{1:t})$  a jednorozměrná pravděpodobnost konkrétní časové trajektorie stavu  $(x_1, \dots, x_t)$ .

Jinak má práce dobrou jazykovou i grafickou úroveň a obsahuje pouze drobné formální nedostatky.

### ***Vyjádření k publikacím autora***

Výsledky doktoranda vztahující se k tématu disertační práce představuje 11 konferenčních příspěvků, většina z nich na kvalitních mezinárodních konferencích. Publikace výsledků v časopise zatím chybí.

Oceňuji také, že jedna publikace autora vznikla v rámci mezinárodní spolupráce s profesorem Anthony Quinnem při pobytu na Trinity College Dublin.

### ***Připomínky oponenta a náměty k odborné diskusi***

1. Definice 1.1, 1.2 a řada následujících "definic", které popisují vlastnosti aproximací metodami MC, by správně měly být nazvány *Lemma*, *Věta* nebo *Tvrzení*.
2. Tvrzení o přesnosti odhadu (1.61) na str. 49 je obráceně.
3. Pro systémy, které vznikají diskretizací spojitých dynamických systému, je třeba definovat model (2.1) jako sdruženou distribuci  $x_{t+1}$  a  $y_t$ , protože šumy procesu a měření mohou být vzájemně korelované. Jak by tato vlastnost stavového modelu ovlivnila aplikovatelnost prezentovaných metod Monte Carlo?
4. Popis časových indexů pod vztahem (2.4) není příliš vhodný, protože neumožňuje formulovat např. úlohu predikce (viz tabulka 2.1). Správná definice pojmu  $l$ -krokový prediktor na str. 53 nevede na hustotu  $p_\theta(x_l | y_{1:t})$ , kde  $l > t$ , ale  $p_\theta(x_{t+l} | y_{1:t})$ .
5. Porovnání RMSE na obr. 2.2 na základě jednoho běhu filtru nemusí být spolehlivým indikátorem pro vyhodnocení kvality jednotlivých filtrů. Postup použitý např. na obr. 4.1 je

průkaznější. Nárůst RMSE ve druhém a třetím sloupci tabulky je částečně způsoben nárůstem amplitudy těchto signálů. Bylo by vhodné tento vliv eliminovat vhodnou normalizací signálů.

6. Vztah (2.20) podle mého názoru definuje podmíněnou pravděpodobnost stavu po datovém kroku

$$\frac{p(y_t | x_t) p(x_t | x_{t-1})}{p(y_t | x_{t-1})} = \frac{p(y_t, x_t | x_{t-1})}{p(y_t | x_{t-1})} = p(x_t | x_{t-1}, y_t)$$

a nikoli návrhovou hustotu.

7. Jde ve vztahu (3.2) opravdu o maticové funkce? Jak je definován skalární součin pro matice?
8. Vzniká problém s degenerací trajektorií stavů řešený v kapitole 4 také při plnohodnotné implementaci zpětného běhu (forward filtering and backward smoothing)? Proč nebyla tato metoda využita alespoň pro porovnání dosažitelných výsledků – je důvodem extrémní výpočetní složitost?
9. Formulace prvního odstavce na str. 105 je dle mého názoru chybná. Zatímco např. šum procesu u Kalmanova filtru je náhodná veličina a je správné při simulaci stochastických systémů generovat v každém časovém okamžiku nové i.i.d. veličiny, subjektivní hustota pravděpodobnosti parametrů popisuje neurčitost parametrů, ale tyto parametry nejsou náhodné, nechovají se jako i.i.d. veličiny (viz např. Peterka, 1981).
10. Formulace vztahu dvojice Kalmanových filtrů na obr. 7.1 (primární a externí filtr) na str. 160 předpokládá existenci dvou filtrů s různými lokálními modely, lokálními pozorováními a lokálními odhady stavu. Vztah (7.7) ale naproti tomu předpokládá, že je možné funkci  $f_e(z_{i,e} | z_{i-1,e}) \Big|_{z_{i,e}=z_i}$  vyhodnotit v bodě  $z_{i,e} = z_i$ . Usuzuji z toho, že oba vektory měření musejí mít stejnou dimenzi, pořadí veličin apod. Můžete vztah obou Kalmanových filtrů v tomto smyslu upřesnit?
11. Použití FPD přístupu k nalezení sdružené hustoty pravděpodobnosti stavu a výstupu primárního filtru (7.4) podmiňováním externí hustotou  $f_e$  ve svém důsledku může způsobit i nárůst neurčitosti odhadu primárního filtru. Bayesovské podmiňování typicky apriorní neurčitost snižuje, nebo ji v případě neinformativní věrohodnostní funkce zachovává. V tomto smyslu je přenos znalosti vždy pozitivní. Umíte chování DT filtru (negativní přenos znalosti) demonstrovat např. experimenty na obr. 7.2 nějak interpretovat?
12. Mám určité pochybnosti o správnosti výsledků prezentovaných na obr. 7.2 až 7.4. Metoda MVF podle [220] je ekvivalentní sériovému nebo paralelnímu zpracování obou vektorů měření v jednom filtru. Vzhledem k tomu, že zvolené kritérium MNSE je vlastně stopa

kovarianční matice odhadu stavu a že Kalmanův filtr splňuje Cramer-Raovu mez (Šimandl et al., 2001), by měl odhad MVF filtrem konvergovat k inverzi Fisherovy informační matice, která je teoreticky dolním limitem dosažitelné hodnoty kovarianční matice odhadu stavu a tedy i její stopy. Přesto některé filtry (iDT filtr) na obr. 7.2 až 7.3 tento limit překračují. Rovněž analytický výpočet ustálené kovarianční matice odhadu stavu pro model poloha-rychlost na str. 167 vede na odlišné hodnoty než výsledky simulací.

### ***Celkové zhodnocení práce***

Ing. Milan Papež pracoval na obtížném tématu a odvedl z mého pohledu nadstandardní množství práce. Seznámil se rozsáhlým publikovaným materiálem, který utřídil a presentoval. Po odstranění nedostatků ovlivňujících srozumitelnost jeho textů by mohla vzniknout kvalitní monografie shrnující vývoj metod MC s dobrým pedagogickým potenciálem.

Práce Ing. Milana Papeže obsahuje také originální výsledky autora, které jsou přínosem pro vědeckou komunitu v oboru. Splňuje proto podle mého názoru nároky na udělení akademického titulu „doktor“ v oboru Kybernetika.

Doporučuji disertační práci Ing. Ing. Milana Papeže k obhajobě.

V Praze 1. 4. 2018



prof. Ing. Vladimír Havlena, CSc.